Reviews

Normal-mode-based Refinement of an F-actin Model Against Fibre Diffraction Data

Yinghao Wu[3] & Jianpeng Ma[1,2,3]

[1] Graduate Program of Structural and Computational Biology and Molecular Biophysics
[2] Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX77030
[3] Department of Bioengineering, Rice University, Houston, TX 77005

Received 17th February 2004; accepted in revised form 16th March 2004

ABSTRACT

This review briefly summarizes the recent development in fibre diffraction refinement of flexible filamentous systems using long-range normal modes as adjustable parameters. Among all the long-range modes, the low-frequency bending modes were found to contribute the most to the improvement of the refinement. The use of several low-frequency modes in the refinement decreased significantly both R- and R_{free} -factors, demonstrating the advantage of this procedure in achieving a good refinement without the risk of over-fitting. Moreover, the study provided strong evidence that substantial errors in conventional refinements are due to long-range deformations, especially the bending, of the filaments. These intrinsic deformations must be properly accounted for in order to improve the refinement efficiency.

Fibre diffraction is widely employed for studying structures of biologically important filamentous systems (Stubbs, 1999) ranging from simple polypeptides to cytoskeletal filaments and filamentous viruses. In fibre diffraction experiments, the fibre specimens are aligned axially, but not azimuthally, and diffraction patterns are cylindrically averaged, which leads to the characteristic layer lines. Because of this averaging, the number of independent diffractions of fibres is significantly smaller than that from crystals. One usually does not have sufficient data to refine the Cartesian coordinates of each atom in the fibres. This feature imposes a severe challenge in choosing proper refinement parameters (Wang and Stubbs, 1993).

In the conventional fibre diffraction refinement procedures, fibres are assumed to be perfect straight helices. In reality, however, fibres are flexible and undergo all sorts of static and dynamic disordering that contribute to the modulation of diffraction patterns. Such disordering should be taken into account in refinement in order to obtain more accurate models.

In recent studies (Ming *et al.* 2003a; Ming *et al.* 2003b), collective long-range deformations of a filament, such as bending, twisting and stretching, were effectively modeled by long-range normal modes of the filament. Following that, a small set of low-frequency modes was used as adjustable parameters to refine the model of *F*-actin (Wu and Ma 2004), a system known to be highly flexible and dynamic (Egelman 2001; Egelman *et al.* 1982; Egelman and Orlova 1995; Galkin *et al.* 2002; Huxley *et al.* 1994; Kojima *et al.* 1994; Orlova *et al.* 2001; Wakabayashi *et al.* 1994). As an approximation, deformations were assumed to occur within the periodic repeat, the helical unit cell, of the *F*-actin filaments. The straight rigid filament model used in conventional refinement procedures was thereby substituted by wave-like conformations described by normal modes.

With the new refinement procedure, it was demonstrated that the long-range deformations of the filaments, especially the bending, are the dominant sources of refinement errors. By taking into account such deformations in the refinement, with only 7~9 low-frequency normal modes as adjustable parameters, it was possible to lower both *R*- and R_{free}-factors by ~2%. Therefore, the long-range filament deformations of fibres indeed need to be properly accounted for in order to improve the refinement.

Structural Models of *F*-actin

The various states of the monomeric G-actin subunit have been visualized by X-ray crystallography (Kabsch *et al.* 1990; McLaughlin *et al.* 1993; Otterbein *et al.* 2001; Robinson *et al.* 1999; Schutt *et al.* 1993). The polymeric *F*-actin filaments have several models. The standard Holmes model (Holmes *et al.* 1990) was established by combining fibre diffraction data and the X-ray structure of G-actin subunit (Kabsch *et al.* 1990). In this model, *F*-actin filaments have two right-handed long-pitch helical strands with a rise of 27.5 Å and a rotation angle of -166.15° per monomer. There are 13 subunits in the minimum repeat of the double-stranded helix (the 13-subunit repeat) with a length of 37.5 nm. A second model (Tirion *et al.* 1995) was refined on the basis of the standard Holmes model, with the deformations of the G-actin subunits modeled by normal modes. Although a reduced number of refinement parameters were used by employing normal modes calculated on a single G-actin sub-unit, no effects of long-range filament deformations were considered. Another model of *F*-actin (Lorenz *et al.* 1993) was obtained using directed mutation algorithm. However, too many parameters were used in the refinement, which has a potential problem of over-fitting (Tirion *et al.* 1995). A new model was provided by Kenneth C. Holmes (unpublished results, personal communication). In this model, each subunit was refined as five rigid bodies, the four G-actin domains and a heptapeptide phalloidin. The *R*-factor of

this model is 8.7% and it was used as the starting model for the new normal-mode-based refinement (Wu and Ma, 2004).

Normal-mode-based Refinement Procedure

In the new refinement procedure, in order to study the effects of long-range deformations on refinement, the four-domains of G-actin were treated as rigid-bodies. Various assemblies of G-actin subunits were used as helical unit cells in the refinement against fibre diffraction data (Holmes *et al.* 1990). Within each unit cell, no helical symmetry was considered so that there was only one asymmetric unit in each unit cell and all atoms were treated as unique. The approximate assumption was that *F*-actin filaments deform in the unit of the helical unit cells. This assumption would result in additional layer lines generated in the calculations.

The low-frequency normal modes of the helical unit cells were determined using Cα-based normal mode analysis, the anisotropic network model (ANM) (Atilgan *et al.* 2001) and the substructure synthesis method (SSM) (Ming *et al.* 2003b). The eigenvectors of normal modes describe the Cα-based trajectories of molecular deformations. The coordinates for all atoms in the unit cell along the deformational trajectories were generated by superimposing the four domains of G-actin subunit, as rigid-bodies, onto the Cα-trajectories along the eigenvectors of normal modes. This assumption of the domains as rigid bodies is reasonable, given that only the low-frequency modes were chosen to describe the long-range collective deformations of the filament in the refinement. Additionally, inter-domain constrains suggested by Holmes in his recent work (unpublished results, personal communication) are also applied (Wu and Ma 2004) to maintain constant distances between two neighboring Cα-atoms located at the boundary of different domains.

Refinement Using Individual Low-frequency Modes of A 13-subunit Repeat

The minimum 13-subunit repeat of *F*-actin was first used as a helical unit cell for the refinement. The low-frequency normal modes of the repeat were calculated (Atilgan *et al.* 2001) and individually employed as adjustable parameters to refine the model. Results in Figure 1 clearly showed that the refinement was improved to various degrees for all the first 20 lowest-frequency modes plotted (the ceiling of the figure is the R-factor of the starting model, 8.7%). The smallest R-factor (7.4%) was achieved by each of the two very lowest-frequency modes (the 1st and 2nd vibrational modes after excluding the first six zero modes) that are two perpendicular bending modes. In the first 11 lowest-frequency modes, all modes are various types of bending modes, with the exception of the 3rd and 6th modes (twisting) and the 9th mode (stretching). All these bending modes reduced the *R*-factor to a larger degree (0.4~1.3%) than non-bending ones (0.2~0.3%). The decrease of R-factor caused by the bending modes becomes progressively smaller as the frequency (or modal index) increases. These results clearly suggest that bending modes of *F*-actin filament, especially the very low-frequency ones, have large impacts on the structure refinement.

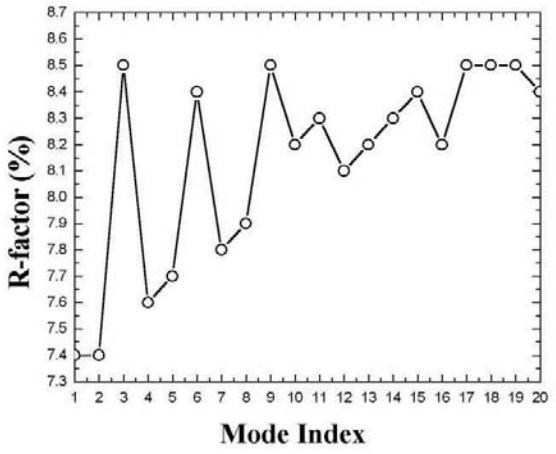

Figure 1. Refinement results using each of the first 20 lowest-frequency modes with 13-subunits as a helical unit cell. The ceiling of the figure is the value of *R*-factor (8.7%) of the starting model of the normal-mode-based refinement (Wu and Ma, 2004).

Refinement Using Individual Low-frequency Bending Modes of Longer Repeats

To investigate the effects of deformations with longer-wavelength, the refinements were further carried out using helical unit cells of longer than 13-subunits. For longer *F*-actin repeats, the low-frequency normal modes were generated by

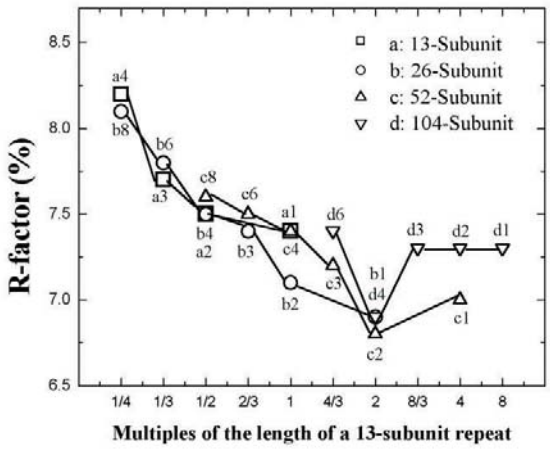

Figure 2. Refinement results using individual bending modes with different lengths of repeating units as helical unit cells. For illustrative purpose, the figure was drawn to align the modes according to their wavelengths. The intervals of the X-axis were expressed as multiples of a single 13-subunit repeat. To make the points in figure spread evenly, all intervals were arranged along the X-axis evenly (rather than based on the numerical values of the actual multiples). The modes were marked according to their half wavelengths by approximately assuming the shapes of waves are trigonometric, *i.e.* those of standing waves. The numerical numbers following the letters at each point were inverted to the types of bending waves in terms of their wavelengths. For example, the 1st type of bending mode of the 13-subunit repeat had a multiple of 1 and was labeled as a1 and the 2nd type of bending mode had a multiple of 1/2 and was labeled as a2, and so on. All the modes for longer repeating units were aligned correspondingly. For clarity, only one of the two degenerate modes was used in each type of wave.

SSM (Ming *et al.* 2003b) based on the modes calculated for a 13-subunit repeat (Ming *et al.* 2003a).

Fig.2 shows the results of refinement by individual bending modes at different lengths of repeats. The X-axis was calibrated to the multiples of the length of a 13-subunit repeat. The curves were aligned according to their approximate half wavelengths that were approximately defined by comparing the shape with that of a standing wave (for the shapes of the first four types of bending modes for standing waves and free vibrations with equal wavelengths, see Fig.3 in reference (Wu and Ma, 2004)). It was noted that there was a specific point at which all the curves converged to a minimum of *R*-factor. The wavelength of that point corresponds to modes with a half wavelength of the length of a 26-subunit repeat (the 1st type of bending mode of a 26-subunit repeat, the 2nd type of bending mode of a 52-subunit repeat, and the 4th type of bending mode of a 104-subunit repeat).

Refinement Using Multiple Modes

The effects of multiple modes on the refinement were also examined because an increased number of degrees of freedom, represented by multiple modes, were expected to better describe the deformations of the filament.

As shown in Fig.3, the use of multiple low-frequency modes indeed further decreased the *R*-factor. For example, the inclusion of 9 lowest-frequency modes of a 13-subunit repeat in the refinement resulted in an *R*-factor of 7.0%. More importantly, the R_{free}-factor was decreased to a similar degree (see Table 1 in reference Wu and Ma, 2004), indicating a real improvement of the refinement. The trends of changes in *R*-factor were similar for larger helical unit cells (26-subunit and 52-subunit repeats). The lowest value of *R*-factor was 6.3% for the 52-subunits using 7~9 lowest-frequency modes.

The Refined Structures

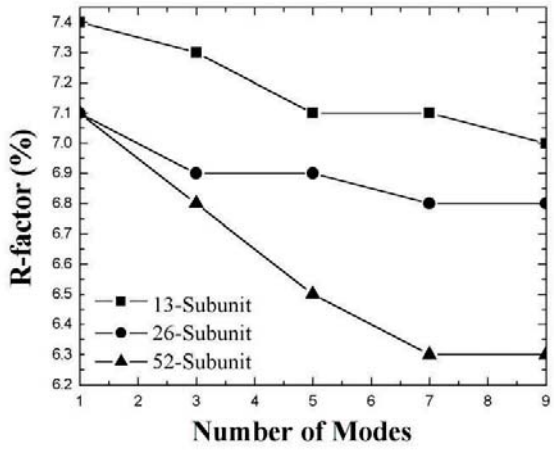

Figure 3. Refinement results using multiple low-frequency normal modes as adjustable parameters with various lengths of repeats as helical unit cells.

The long-range normal-mode refinement procedure improves the fit between the model and the diffraction data mainly by a more realistic modelling of the long-range deformations. Lo-cal structural adjustments in individual subunits and domains are relatively small (see Fig.5 in reference Wu and Ma, 2004). This is illustrated in Fig.4 in which the G-actin subunit in the standard Holmes model is superimposed with the G-actin structure refined using a 52-subunit as a helical unit cell by employing 7~9 lowest-frequency modes (with *R*-factor of 6.3%). The refined G-actin structure is very similar to the new Holmes model that was used as a starting model of this study, with an average root-mean-square (rms) deviation of 0.1~0.2 Å.

Summary

This review briefly summarizes the results of a recent study (Wu and Ma, 2004) on the refinement of the *F*-actin model against fibre diffraction data using long-range normal modes as adjustable parameters to account for the collective filament deformations. In the refinement, each of the four domains of a G-actin subunit was treated as rigid-body in order to separate the contributions of long-range deformations from those of local structural deformations. Among all the low-frequency long-range modes, bending modes were found to contribute the most to the improvement of the refinement. Moreover, the largest decrease of both *R*- and R_{free}- factors was achieved by using multiple low-frequency modes. The results clearly demonstrated the advantage of normal-modes-based refinement in using a small number of adjustable parameters to achieve a good fitting efficiency, which helps avoid the risk of the over-fitting problem. More importantly, the results provided strong evidence that, for fibre diffraction refinement of flexible systems, substantial errors arise from the ignorance of long-range deformations, especially bending, of the filaments. Therefore, the effects of these deformations must be included in the refinement for flexible systems.

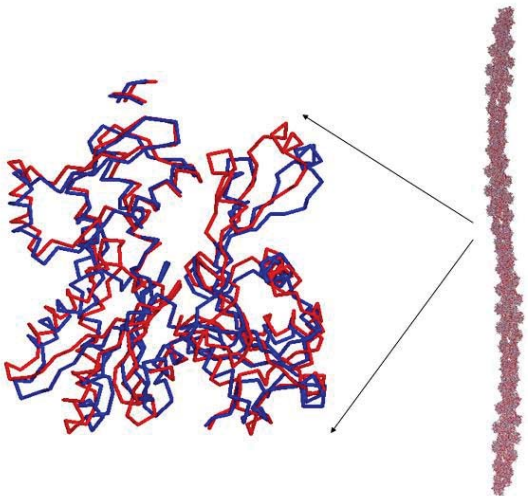

Figure 4. The superposition of the G-actin subunit in the standard Holmes model (blue) with the G-actin structure refined using a 52-subunit as a helical unit cell by employing 7~9 lowest-frequency normal modes (red). The refined G-actin structure is very similar to the new Holmes model that was used as a starting model of this study, with an average rms deviation of 0.1~0.2 Å. The starting model is omitted in the figure for illustration clarity.

In the current normal-mode-based refinement procedure (Wu and Ma, 2004), the contributions from slightly higher-frequency modes with wavelengths similar to the physical sizes of the domains were eliminated due to the use of rigid body for each of the four domains of the G-actin subunit. It is expected to further improve the refinement once such a restriction is lifted to allow local structures to adjust within each individual domain. Another important aspect of the new refinement procedure is that there is only one asymmetric unit in each unit cell in which all atoms were regarded unique and participated in the calculations. This treatment will be especially useful in the later stage of fibre diffraction refinement, when it is possible to determine long-range normal modes from a model refined using conventional methods. However, it needs to be emphasized that, as the size of unit cell used in normal-mode-based refinement increases, the computational cost quickly increases, which may affect the application of the method to systems with very large unit cells. Future development should be devoted to resolve this issue.

Acknowledgements

The authors gratefully acknowledge Professor Kenneth C. Holmes for providing the experimental diffraction data and the unpublished *F*-actin model. J.M. thanks the support from the American Heart Association (AHA-TX0160107Y), the Robert A. Welch Foundation (Q-1512), the National Institutes of Health (R01-GM067801), and the National Science Foundation Career Award (MCB-0237796). J.M. is also a recipient of the Award for Distinguished Young Scholars Abroad from National Natural Science Foundation of China.

References

[1] Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I. (2001) Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys J,* **80(1)**, 505-515.

[2] Egelman EH. (2001) Actin allostery again? *Nat Struct Biol,* **8(9)**, 735-736.

[3] Egelman EH, Francis N, DeRosier DJ. (1982) *F*-actin is a helix with a random variable twist. *Nature,* **298**, 131-135.

[4] Egelman EH, Orlova A. (1995) New insights into actin filament dynamics. *Curr Opin Struct Biol,* **5(2)**, 172-180.

[5] Galkin VE, VanLoock MS, Orlova A, Egelman EH. (2002) A new internal mode in *F*-actin helps explain the remarkable evolutionary conservation of actin's sequence and structure. *Curr Biol,* **12(7)**, 570-575.

[6] Holmes KC, Popp D, Gebhard W, Kabsch W. (1990) Atomic model of the actin filament. *Nature,* **347(6288)**, 44-49.

[7] Huxley HE, Stewart A, Sosa H, Irving T. (1994) X-ray diffraction measurements of the extensibility of actin and myosin filaments in contracting muscle. *Biophys J,* **67**, 2411-2421.

[8] Kabsch W, Mannherz HG, Suck D, Pai EF, Holmes KC. (1990) Atomic structure of the actin:DNase I complex. *Nature,* **347(6288)**, 37-44.

[9] Kojima H, Ishijima A, Yanagida T. (1994) Direct measurement of stiffness of single actin filaments with and without tropomyosin by in vitro nanomanipulation. *Proc Natl Acad Sci USA,* **91(26)**, 12962-12966.

[10] Lorenz M, Popp D, Holmes KC. (1993) Refinement of the *F*-actin model against X-ray fibre diffraction data by the use of a directed mutation algorithm. *J Mol Biol,* **234(3)**, 826-836.

[11] McLaughlin PJ, Gooch JT, Mannherz HG, Weeds AG. (1993) Structure of gelsolin segment 1-actin complex and the mechanism of filament severing. *Nature,* **364(6439)**, 685-692.

[12] Ming D, Kong Y, Wu Y, Ma J. (2003a) Simulation of *F*-actin Filaments of Several Microns. *Biophys J,* **85**, 27-35.

[13] Ming D, Kong Y, Wu Y, Ma J. (2003b) Substructure Synthesis Method for Simulating Large Molecular Complexes. *Proc. Nalt. Acad. Sci. USA,* **100**, 104-109.

[14] Orlova A, Galkin VE, VanLoock MS, Kim E, Shvetsov A, Reisler E, Egelman EH. (2001) Probing the structure of *F*-actin: cross-links constrain atomic models and modify actin dynamics. *J Mol Biol,* **312(1)**, 95-106.

[15] Otterbein LR, Graceffa P, Dominguez R. (2001) The crystal structure of uncomplexed actin in the ADP state. *Science,* **293(5530)**, 708-711.

[16] Robinson RC, Mejillano M, Le VP, Burtnick LD, Yin HL, Choe S. (1999) Domain movement in gelsolin: a calcium-activated switch. *Science,* **286(5446)**, 1939-1942.

[17] Schutt CE, Myslik JC, Rozycki MD, Goonesekere NC, Lindberg U. (1993) The structure of crystalline profilin-beta-actin. *Nature,* **365(6449)**, 810-816.

[18] Stubbs G. (1999) Development in Fibre Diffraction. *Curr. Opin. Struct. Biol.,* **9**, 615-619.

[19] Tirion MM, ben-Avraham D, Lorenz M, Holmes KC. (1995) Normal modes as refinement parameters for the *F*-actin model. *Biophys J,* **68(1)**, 5-12.

[20] Wakabayashi K, Sugimoto Y, Tanaka H, Ueno Y, Takezawa Y, Amemiya Y. (1994) X-ray diffraction evidence for the extensibility of actin and myosin filaments during muscle contraction. *Biophys J,* **67**, 2422-2435.

[21] Wang H, Stubbs G. (1993) Molecular dynamics in refinement against fibre diffraction data. *Acta Crystallogr A,* **49**, 504-513.

[22] Wu Y, Ma J. (2004) Refinement of *F*-actin Model Against Fibre Diffraction Data by Long-range Normal Modes. *Biophys. J.,* **86**, 116-124.